



Implementasi Algoritma K-nearest Neighbor dan Principal Component Analysis untuk Klasifikasi Tingkat Kematangan Ceri Kopi Robusta Berdasarkan Warna

Ayu Wulandari*, Rudi Heriansyah, Lastri Widya Astuti

Department of Informatic Engineering, Indo Global Mandiri University, Indonesia

DOI:

<https://doi.org/10.53697/jkomitek.v5i2.3048>

*Correspondence: Ayu Wulandari

Email: 2023110128@students.uigm.ac.id

Received: 02-10-2025

Accepted: 20-11-2025

Published: 19-12-2025



Copyright: © 2025 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: This study addresses the need for objective and efficient classification of Robusta coffee cherry ripeness, which is crucial for ensuring optimal coffee quality. The research aims to develop an automatic classification system using K-Nearest Neighbor (KNN) and Principal Component Analysis (PCA) based on digital image color features. Employing a quantitative experimental approach, the study utilized 150 digital images of Robusta coffee cherries, categorized into three ripeness levels: unripe, semi-ripe, and ripe. Data were collected using a smartphone camera under controlled lighting, and processed with MATLAB for feature extraction and analysis. The RGB color features were reduced using PCA, and classification was performed with KNN. Evaluation metrics included confusion matrix, accuracy, precision, and recall. The results showed that the system achieved an accuracy of 93.33%, successfully classifying 28 out of 30 test images correctly. These findings indicate that the combination of KNN and PCA provides a reliable and practical solution for automated coffee cherry ripeness classification. The study concludes that this approach can enhance post-harvest processes and support consistent quality in Robusta coffee production.

Keywords: Classification, Coffee Cherry, Digital Image, K-Nearest Neighbor, Principal Component Analysis

Pendahuluan

Kopi merupakan salah satu komoditas perkebunan yang memiliki peran penting dalam perekonomian Indonesia, menempati posisi keempat dunia dalam hal produksi dan ekspor setelah Brasil, Vietnam, dan Kolombia (Rahardjo et al., 2020; Novita et al., 2023). Salah satu varietas utama yang banyak dibudidayakan adalah kopi Robusta (*Coffea canephora*), yang dikenal dengan kandungan kafein tinggi dan karakteristik rasa yang khas (Widawati et al., 2023; Kurnia, 2023). Daerah Sumatera Selatan, khususnya kota Pagar Alam, menjadi salah satu sentra produksi kopi Robusta yang signifikan di Indonesia (Gusmaliza & Aminah, 2024; Novita et al., 2023). Tingkat kematangan ceri kopi Robusta sangat menentukan kualitas akhir biji kopi, di mana perubahan warna kulit ceri dari hijau ke merah menandakan proses pematangan yang memengaruhi aroma dan cita rasa kopi (Alam et al., 2023; Widawati et al., 2023).

Penentuan tingkat kematangan ceri kopi Robusta umumnya masih dilakukan secara manual oleh petani, sehingga berpotensi menimbulkan ketidakkonsistenan dalam pemilahan dan penurunan mutu hasil panen (Rioarda Irfa'i et al., 2021; Alam et al., 2023). Proses klasifikasi kematangan yang akurat sangat penting untuk memastikan kualitas kopi yang optimal, namun metode konvensional seringkali kurang efisien dan subjektif (Adenugraha et al., 2022; Sari, 2023). Penelitian sebelumnya telah mencoba mengatasi permasalahan ini dengan berbagai pendekatan, seperti penggunaan algoritma Fuzzy C-Means yang menghasilkan akurasi klasifikasi sebesar 75% (Rioarda Irfa'i et al., 2021) dan kombinasi K-Nearest Neighbor (KNN) dengan Principal Component Analysis (PCA) pada klasifikasi buah pisang yang mencapai akurasi 90,9% (Adenugraha et al., 2022).

Meskipun telah ada upaya untuk mengotomatisasi klasifikasi tingkat kematangan buah menggunakan citra digital, tantangan utama yang dihadapi adalah tingginya dimensi data warna yang dihasilkan dari citra, serta kebutuhan akan metode klasifikasi yang sederhana namun efektif (Sari, 2023; Adenugraha et al., 2022). KNN dikenal sebagai algoritma yang sederhana dan efektif dalam menangani klasifikasi berbasis fitur warna, sedangkan PCA berfungsi untuk mereduksi dimensi data tanpa kehilangan informasi penting (Sari, 2023; Adenugraha et al., 2022). Kombinasi kedua metode ini telah terbukti meningkatkan akurasi dan efisiensi dalam beberapa penelitian terkait klasifikasi buah berbasis citra (Adenugraha et al., 2022; Sari, 2023).

Namun, penelitian mengenai implementasi kombinasi KNN dan PCA untuk klasifikasi tingkat kematangan ceri kopi Robusta berdasarkan warna masih sangat terbatas, terutama pada penggunaan data citra digital yang diambil dalam kondisi pencahayaan terkontrol dan jumlah data yang memadai (Rioarda Irfa'i et al., 2021; Sari, 2023). Selain itu, belum banyak penelitian yang secara khusus memfokuskan pada pengembangan sistem klasifikasi otomatis berbasis citra digital untuk ceri kopi Robusta di Indonesia, sehingga diperlukan kajian lebih lanjut untuk mengisi kesenjangan tersebut (Adenugraha et al., 2022; Gusmaliza & Aminah, 2024).

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi otomatis tingkat kematangan ceri kopi Robusta berbasis citra digital dengan mengimplementasikan algoritma K-Nearest Neighbor (KNN) yang dioptimalkan melalui Principal Component Analysis (PCA). Urgensi penelitian ini terletak pada upaya meningkatkan akurasi dan efisiensi proses klasifikasi kematangan ceri kopi, yang pada akhirnya dapat mendukung peningkatan mutu hasil panen kopi Robusta di Indonesia. Kebaruan penelitian ini terletak pada penerapan kombinasi KNN dan PCA pada data citra digital ceri kopi Robusta dengan jumlah data yang lebih besar dan kondisi pengambilan gambar yang terstandarisasi, sehingga diharapkan dapat memberikan kontribusi signifikan dalam pengembangan teknologi klasifikasi buah berbasis citra di bidang pertanian (Adenugraha et al., 2022; Sari, 2023).

Metodologi

Jenis dan Metode Penelitian

Penelitian ini merupakan penelitian terapan dengan pendekatan kuantitatif yang bertujuan untuk mengembangkan sistem klasifikasi otomatis tingkat kematangan ceri kopi Robusta berbasis citra digital menggunakan algoritma K-Nearest Neighbor (KNN) dan Principal Component Analysis (PCA). Metode penelitian yang digunakan adalah eksperimen laboratorium, di mana data citra ceri kopi Robusta dikumpulkan, diproses, dan dianalisis secara sistematis untuk menguji efektivitas kombinasi KNN dan PCA dalam klasifikasi tingkat kematangan berdasarkan fitur warna citra (Adenugraha et al., 2022; Mujidah & Agustin, 2024). Penelitian ini mengacu pada tahapan penelitian kuantitatif sebagaimana dijelaskan oleh Sugiyono (2022) dan Cresswell (2022), yang menekankan pada proses pengumpulan data, analisis data, serta pengujian hipotesis secara objektif dan terukur.

Instrumen dan Teknik Analisis Data

Instrumen utama dalam penelitian ini adalah perangkat keras berupa laptop dengan spesifikasi Intel Core i5 dan RAM 8GB, smartphone iPhone 13 untuk pengambilan citra, lampu LED sebagai sumber pencahayaan, tripod, serta perangkat lunak MATLAB untuk pengolahan citra dan implementasi algoritma. Teknik analisis data yang digunakan meliputi preprocessing citra (penghapusan latar belakang, penyesuaian dimensi, dan pengurangan noise), ekstraksi fitur warna RGB, reduksi dimensi menggunakan PCA, serta klasifikasi menggunakan KNN. Evaluasi kinerja sistem dilakukan dengan menggunakan metrik confusion matrix, akurasi, presisi, dan recall (Mujidah & Agustin, 2024; Emzir, 2021). Prosedur analisis data ini sejalan dengan pendekatan yang digunakan dalam penelitian klasifikasi kualitas biji kopi berbasis citra digital (Mahadany, 2024; Angriani, 2023).

Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh ceri kopi Robusta yang terdapat di wilayah penelitian, sedangkan sampel yang digunakan berupa 150 citra digital ceri kopi Robusta yang mewakili tiga kelas tingkat kematangan: mentah, setengah matang, dan matang. Sampel citra diambil secara purposive sampling dengan mempertimbangkan variasi tingkat kematangan dan kondisi pencahayaan yang seragam. Pada tahap awal, digunakan dataset sementara sebanyak 24 citra untuk pengujian awal metode, yang kemudian diperluas menjadi 150 citra untuk pengujian utama (Adenugraha et al., 2022; Mujidah & Agustin, 2024). Teknik pengambilan sampel ini mengacu pada prinsip-prinsip sampling dalam penelitian kuantitatif sebagaimana dijelaskan oleh Sudaryono (2022) dan Cresswell (2022).

Prosedur Penelitian

Prosedur penelitian diawali dengan pengumpulan data citra ceri kopi Robusta menggunakan smartphone dengan latar belakang putih dan pencahayaan lampu LED untuk memastikan konsistensi warna. Selanjutnya, citra yang diperoleh diproses melalui

tahapan preprocessing untuk menghilangkan noise dan menyesuaikan dimensi citra. Setelah itu, dilakukan ekstraksi fitur warna RGB dari setiap citra, yang kemudian direduksi dimensinya menggunakan PCA untuk memperoleh fitur yang paling relevan. Fitur hasil reduksi selanjutnya digunakan sebagai input pada algoritma KNN untuk proses klasifikasi tingkat kematangan ceri kopi. Hasil klasifikasi dievaluasi menggunakan confusion matrix, akurasi, presisi, dan recall untuk menilai kinerja sistem (Adenugraha et al., 2022; Mahadany, 2024; Sugiyono, 2022). Prosedur ini dirancang secara sistematis dan logis sesuai dengan standar penelitian kuantitatif dan pengembangan sistem klasifikasi berbasis citra digital (Cresswell, 2022; Emzir, 2021).

Hasil dan Pembahasan

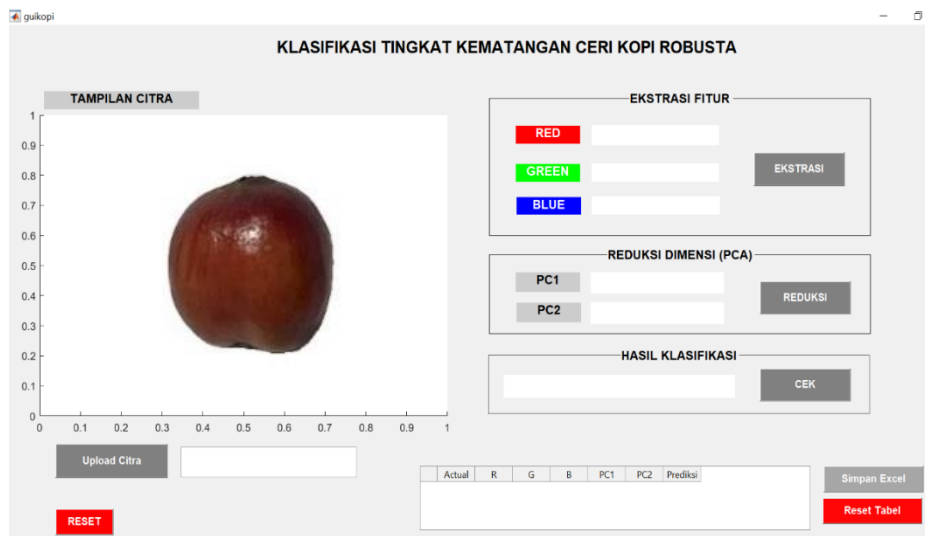
Penelitian ini menggunakan 150 citra ceri kopi Robusta yang diperoleh dari kebun kopi di Kelurahan Jokoh, Kota Pagar Alam. Proses pengambilan citra dilakukan di dalam ruangan dengan pencahayaan tambahan berupa lampu LED untuk memastikan kualitas gambar yang seragam. Dari total data, 80% digunakan sebagai data latih (40 citra per kelas: mentah, setengah matang, matang) dan 20% sebagai data uji (10 citra per kelas). Pembagian ini bertujuan untuk menjaga keseimbangan data serta memastikan evaluasi performa sistem berjalan optimal.



Gambar 1. Pohon Kopi Robusta di Kelurahan Jokoh, Kota Pagar Alam

Tampilan Antarmuka Aplikasi

Aplikasi klasifikasi dikembangkan menggunakan MATLAB GUI dengan rancangan antarmuka yang interaktif dan mudah digunakan. Fitur utama meliputi unggah citra, ekstraksi nilai RGB, reduksi PCA, klasifikasi menggunakan KNN, penyimpanan hasil ke Excel, serta reset citra maupun tabel. Komponen antarmuka terdiri dari tampilan citra (axes), panel kontrol, kolom nilai RGB, kolom nilai PCA (PC1, PC2), kolom hasil klasifikasi, serta tabel hasil klasifikasi



Gambar 1. Tampilan Antarmuka Aplikasi

Aplikasi klasifikasi kematangan ceri kopi Robusta berbasis MATLAB GUI memiliki alur penggunaan yang mudah dipahami. Pertama, pengguna mengunggah citra kopi melalui tombol Upload Image, dan citra akan ditampilkan pada layar. Selanjutnya, tombol Ekstrak RGB digunakan untuk menghitung nilai rata-rata komponen warna merah, hijau, dan biru, yang kemudian direduksi dimensinya melalui Proses PCA menjadi dua komponen utama, PC1 dan PC2. Dengan menekan tombol Klasifikasi, sistem menjalankan algoritma KNN untuk menentukan tingkat kematangan kopi, yaitu mentah, setengah matang, atau matang, yang hasilnya ditampilkan pada kolom keluaran serta tercatat dalam tabel. Pengguna juga dapat menyimpan hasil tersebut ke dalam file Excel melalui tombol Simpan, atau menghapus data dengan menekan tombol Reset agar aplikasi siap digunakan untuk citra selanjutnya.

Setelah dilakukan klasifikasi menggunakan antarmuka diatas, lalu didapatkan hasil pengklasifikasian ceri kopi Robusta sesuai dengan tingkat kematangannya. Dari total 30 data uji yang sudah dilakukan pengujian, didapatkan hasil prediksi sebagai berikut:

Tabel 1. Tabel hasil klasifikasi KNN pada data uji ceri kopi Robusta

Actual	R	G	B	PC1	PC2	Prediksi
ujimentah1	200,371	200,962	182,441	-1,25	-0,62	mentah
ujimentah2	197,749	199,915	181,521	-1,62	-0,62	mentah
ujimentah3	206,738	209,562	188,828	0,42	-0,64	setengahmatang
ujimentah4	199,303	203,028	182,317	-1,2	-0,8	mentah
ujimentah5	197,2	199,189	180,609	-1,79	-0,66	mentah
ujimentah6	192,937	195,953	179,337	-2,48	-0,5	mentah
ujimentah7	198,535	200,914	184,255	-1,27	-0,39	mentah
ujimentah8	201,132	203,067	185,948	-0,76	-0,39	mentah
ujimentah9	204,678	206,799	186,847	-0,11	-0,63	mentah
ujimentah10	195,084	196,374	178,605	-2,33	-0,63	mentah

Actual	R	G	B	PC1	PC2	Prediksi
ujisetengah1	222,671	217,57	201,226	3,32	-0,02	setengahmatang
ujisetengah2	216,578	208,33	187,574	1,08	-0,8	setengahmatang
ujisetengah3	212,614	203,639	179,207	-0,25	-1,36	setengahmatang
ujisetengah4	225,812	214,684	193,15	2,75	-0,76	setengahmatang
ujisetengah5	222,758	212,872	191,712	2,26	-0,75	setengahmatang
ujisetengah6	219,786	215,073	194,933	2,41	-0,51	setengahmatang
ujisetengah7	215,135	205,817	182,571	0,39	-1,18	setengahmatang
ujisetengah8	214,097	204,183	187,955	0,6	-0,4	mentah
ujisetengah9	212,631	204,156	184,431	0,19	-0,79	setengahmatang
ujisetengah10	215,833	210,224	189,607	1,32	-0,7	setengahmatang
ujimatang1	196,211	181,045	176,86	-3,45	0,34	matang
ujimatang2	201,718	181,715	181,359	-2,59	0,74	matang
ujimatang3	194,437	182,055	183,685	-3	1,06	matang
ujimatang4	194,117	183,587	183,321	-2,95	0,91	matang
ujimatang5	199,499	184,136	183,18	-2,47	0,79	matang
ujimatang6	206,451	193,492	194,892	-0,3	1,35	matang
ujimatang7	212,11	197,188	196,22	0,57	1,14	matang
ujimatang8	204,804	190,99	190,783	-0,94	1,09	matang
ujimatang9	202,689	187,109	188,291	-1,59	1,12	matang
ujimatang10	209,811	201,275	200,979	1,03	1,4	matang

Berdasarkan hasil yang ditampilkan pada Tabel 1, dari total 30 data uji ceri kopi Robusta, sebanyak 28 data berhasil diprediksi dan diklasifikasikan dengan benar sesuai tingkat kematangannya. Hasil ini menunjukkan bahwa sistem memiliki tingkat akurasi yang cukup tinggi dalam mengenali kelas kematangan ceri kopi.

Evaluasi Kinerja Klasifikasi

Evaluasi kinerja sistem klasifikasi sangat penting untuk mengukur seberapa efektif dan akurat model yang dibangun dalam mengelompokkan data ke dalam kelas yang sesuai. Dalam penelitian ini, evaluasi dilakukan terhadap hasil klasifikasi algoritma KNN yang telah diterapkan pada data uji citra buah kopi Robusta. Metode evaluasi yang digunakan dalam penelitian ini mencakup perhitungan *confusion matrix*, serta pengukuran metrik evaluasi seperti akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score* untuk masing-masing kelas yaitu mentah, setengah matang, dan matang.

```

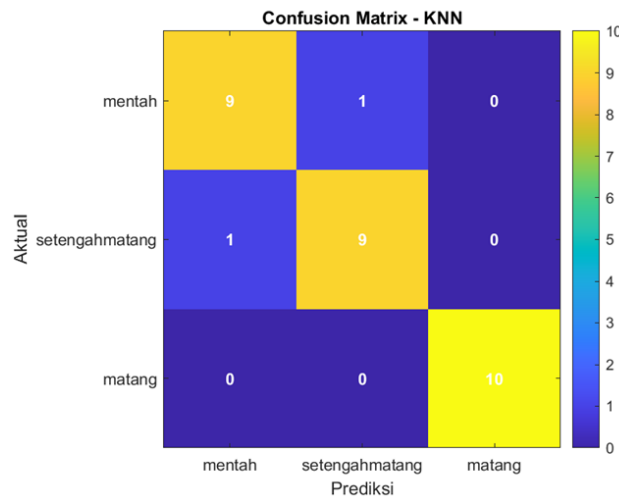
Command Window
>> evaluasi_kinerja('hasil_klasifikasi_20250703_011518.xlsx')

Confusion Matrix:

Evaluasi Kinerja per Kelas:
Kelas      TP  FP  FN  Precision  Recall  F1Score
mentah      9   1   1    0.90     0.90    0.90
setengahmatang  9   1   1    0.90     0.90    0.90
matang     10   0   0    1.00     1.00    1.00

Akurasi Total: 93.33%
    
```

Gambar 3. Hasil Evaluasi Kinerja Klasifikasi



Gambar 4. Hasil Evaluasi Kinerja Klasifikasi

Secara keseluruhan yang dapat dilihat pada Gambar 4 dan 5, proses klasifikasi yang dilakukan dengan metode algoritma KNN dengan parameter $k=3$ dalam membedakan tiga kelas kematangan ceri kopi Robusta yaitu matang, setengah matang, dan matang. Dari 30 citra data uji yang diuji secara satu-persatu pada GUI Matlab atau pada antarmuka aplikasi, sebanyak 28 citra berhasil diklasifikasikan dengan benar dengan tingkat akurasi keseluruhan mencapai 93,33%.

Hasil dari evaluasi kinerja klasifikasi lebih lanjut memperlihatkan performa sistem, dimana nilai yang didapatkan pada hasil evaluasi kinerja untuk kelas mentah dan setengah matang yang memiliki persamaan hasil yaitu nilai presisi (*precision*) mencapai nilai 0.90, nilai *recall* mencapai nilai 0.90, dan nilai F1-score mencapai 0.90. Lalu, untuk kelas matang mendapatkan nilai 1.00 untuk seluruh hasil evaluasi kinerja. Dapat dilihat dari hasil evaluasi kinerja yang didapatkan bahwa mengindikasikan performa klasifikasi yang sangat seimbang antara presisi dan sensitivitas.

Dengan demikian, pendekatan kombinasi antara ekstraksi fitur warna RGB, reduksi PCA, dan klasifikasi dengan metode KNN dengan $k = 3$ memberikan hasil yang presisi dan

mudah diimplementasikan terhadap pola data yang telah dilatih dan terbukti mampu menghasilkan sistem klasifikasi yang sederhana dan unggul.

Simpulan

Kesimpulan dari penelitian ini menunjukkan bahwa implementasi kombinasi algoritma K-Nearest Neighbor (KNN) dan Principal Component Analysis (PCA) pada klasifikasi tingkat kematangan ceri kopi Robusta berbasis citra digital mampu menghasilkan sistem klasifikasi yang akurat dan efisien. Dengan menggunakan fitur warna RGB yang direduksi melalui PCA, sistem berhasil mengklasifikasikan 28 dari 30 data uji dengan tingkat akurasi mencapai 93,33%. Evaluasi kinerja menggunakan confusion matrix, presisi, recall, dan F1-score memperlihatkan performa yang seimbang dan konsisten pada setiap kelas kematangan, khususnya pada kelas matang yang memperoleh nilai evaluasi sempurna. Temuan ini menegaskan bahwa pendekatan yang digunakan dapat diandalkan untuk mendukung proses pemilahan ceri kopi secara otomatis dan konsisten, sehingga berpotensi meningkatkan mutu hasil panen kopi Robusta di Indonesia.

Namun, penelitian ini memiliki beberapa keterbatasan, di antaranya jumlah data uji yang masih terbatas dan pengambilan citra yang hanya dilakukan dalam satu kondisi pencahayaan. Selain itu, sistem belum diuji pada variasi lingkungan dan perangkat pengambilan gambar yang berbeda, sehingga generalisasi hasil masih perlu dikaji lebih lanjut. Untuk penelitian selanjutnya, disarankan agar dilakukan pengujian dengan jumlah data yang lebih besar, variasi kondisi pencahayaan, serta pengembangan algoritma klasifikasi yang lebih adaptif terhadap perubahan lingkungan. Implikasi praktis dari penelitian ini adalah tersedianya sistem klasifikasi otomatis yang dapat diintegrasikan dalam proses pascapanen kopi, sehingga membantu petani dan pelaku industri dalam meningkatkan efisiensi dan konsistensi kualitas kopi Robusta secara berkelanjutan.

Referensi

- Adenugraha, S. P., Arinal, V., & Mulyana, D. I. (2022). Klasifikasi kematangan buah pisang ambon menggunakan metode KNN dan PCA berdasarkan citra RGB dan HSV. *Jurnal Media Informatika Budidarma*, 6(1), 9–17. <https://doi.org/10.30865/mib.v6i1.3287>
- Alam, I., Warkoyo, W., & Siskawardani, D. D. (2023). Karakteristik tingkat kematangan buah kopi robusta (*Coffea canephora* A. Froehner) dan buah kopi arabika (*Coffea arabica* Linnaeus) terhadap mutu dan cita rasa seduhan kopi. *Food Technology and Halal Science Journal*, 5(2), 169–185. <https://doi.org/10.22219/fths.v5i2.21925>
- Cresswell, J. W. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Emzir. (2021). *Metodologi penelitian kualitatif dan kuantitatif*. Rajawali Pers.

- Gusmaliza, D., & Aminah, S. (2024). Sistem identifikasi kualitas biji kopi robusta berbasis image processing dengan support vector machine. *Edumatic: Jurnal Pendidikan Informatika*, 8(2), 744–753. <https://doi.org/10.29408/edumatic.v8i2.28008>
- Kurnia, I. G. A. M. (2023, April 20). Perbedaan mendasar kopi arabika dan kopi robusta. Dinas Pertanian Kabupaten Buleleng. <https://distan.bulelengkab.go.id/informasi/detail/artikel/perbedaan-mendasar-kopi-arabika-dan-kopi-robusta-56>
- Mahadany, S. A. (2024). Desain klasifikasi cherri kopi menggunakan metode k-nearest neighbor. *Jurnal Ilmu Teknologi dan Pertanian*, 5(1), 1–12. <https://journal.upy.ac.id/index.php/jitap/article/view/7410/4319>
- Mujidah, S., & Agustin, R. (2024). Analisis performa algoritma K-nearest neighbor dan reduksi dimensi principal component analysis pada klasifikasi data. *Jurnal Ilmu Komputer dan Informatika*, 4(1), 61–72. <https://doi.org/10.54082/jiki.168>
- Novita, D., Sari, D. R. P., & Pratiwi, R. (2023). Analisis produksi dan ekspor kopi Indonesia. *Jurnal Agribisnis Indonesia*, 11(1), 45–56. <https://doi.org/10.29244/jai.2023.11.1.45-56>
- Rahardjo, S., Prasetyo, B., & Santoso, D. (2020). Analisis ekspor kopi Indonesia ke pasar dunia. *Jurnal Ekonomi dan Kebijakan*, 13(2), 123–134. <https://doi.org/10.24843/JEK.2020.v13.i02.p03>
- Rioarda Irfa'i, M., Fatkhurrozi, B., & Setyowati, I. (2021). Klasifikasi tingkat kematangan buah kopi menggunakan algoritma fuzzy c-means. *THETA OMEGA: Journal of Electrical Engineering*, 1(2), 45–52.
- Sari, D. R. P. (2023). Metode principal component analysis (PCA) sebagai penanganan asumsi multikolinearitas. *Parameter: Jurnal Matematika, Statistika dan Terapannya*, 2(2), 115–124. <https://doi.org/10.30598/parameter.v2i02pp115-124>
- Sudaryono. (2022). *Metode penelitian kuantitatif*. Bumi Aksara.
- Sugiyono. (2022). *Metode penelitian kuantitatif, kualitatif, dan R&D* (2nd ed.). Alfabeta.
- Widawati, L., Moulina, M. A., & Rikardo, R. (2023). Karakteristik mutu sirup kopi robusta (*Coffea canephora*) dan sirup kopi arabika (*Coffea arabica*) dengan penambahan konsentrasi gula. *SINTA Journal (Science, Technology, and Agricultural)*, 4(1), 1–8. <https://doi.org/10.37638/sinta.4.1.1-8>